

m m

02-80
04-20
03-00

#4

OIPE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/001,851

DATE: 12/11/2001

TIME: 11:57:51

Input Set : A:\10147561.app

Output Set: N:\CRF3\12112001\I001851.raw

ENTERED

3 <110> APPLICANT: MEYERS, Rachel A.
 4 WILLIAMSON, Mark
 6 <120> TITLE OF INVENTION: 47169 and 33935, Novel Human Glycosyl Transferases and
 7 Uses Thereof
 9 <130> FILE REFERENCE: 10147-56U1
 C--> 11 <140> CURRENT APPLICATION NUMBER: US/10/001,851
 12 <141> CURRENT FILING DATE: 2001-11-20
 14 <150> PRIOR APPLICATION NUMBER: US 60/249,939
 15 <151> PRIOR FILING DATE: 2000-11-20
 17 <160> NUMBER OF SEQ ID NOS: 29
 19 <170> SOFTWARE: PatentIn Ver. 2.1
 21 <210> SEQ ID NO: 1
 22 <211> LENGTH: 3985
 23 <212> TYPE: DNA
 24 <213> ORGANISM: Homo sapiens
 26 <400> SEQUENCE: 1
 27 ggctgagctg ctgccgccgc cgggcggacg ggccggaacg cggagctggg ggcggcgcg 60
 28 cggggccggc ggggcgcggc ggggctgacc ggcccgatg aggcggaagg agaagcggct 120
 29 cctgcaggcg gtggcgctgg tgctggcggc cctggctctc ctgcccaacg tggggctttg 180
 30 ggcgctgtac cgcgagcggc agcccgacgg caccocctgg ggatcggggg cggcgggtggc 240
 31 gccggcgggc ggacagggct cacacagtgc acaaaagaaa acgtttttct tgggagatgg 300
 32 gcagaagctg aaggactggc atgacaagga ggccatccgg agggacgctc agcgcgtagg 360
 33 aaatggagaa caaggaagac cttaccccat gaccgatgct gagagagtgg atcaggcata 420
 34 ccgagaaaaa ggattttaaca tctacgtcag tgataaaatc tccttgaatc gctctctccc 480
 35 agatatccgg caccaaaact gcaacagcaa gcgctacctg gagacacttc ccaacacaa 540
 36 catcatcatc cccttccaca acgagggctg gtcctccctc ctccgcaccg tccacagtgt 600
 37 gctcaatcgc tcgcctccag agctggctgc cgagattgta ctggtcgacg acttcagtga 660
 38 tcgagagcac ctgaagaagc ctcttgaaga ctacatggcc cttttcccca gtgtgaggat 720
 39 tcttcgaacc aagaaacggg aaggcgctgat aaggaccgga atgctggggg cctcagtggt 780
 40 aactggggat gtcattcacat tcttggttcc acactgtgaa gccaatgtca actggcttcc 840
 41 ccccttgctt gaccgcattg ctccggaaccg caagaccatt gtgtgcccga tgattgatgt 900
 42 aattgaccat gacgactttc ggtacgagac acaggcaggg gatgccatgc ggggagcctt 960
 43 tgactgggag atgtactaca agcggatccc gatccctcca gaactgcaga aagctgaccc 1020
 44 cagcgaccca tttgagtctc ccgtgatggc cgggtggactg ttccgcgtgg atcgggaagt 1080
 45 gttctgggaa ctccggcggt atgaccaggg cttggagatc tggggagggg agcagtatga 1140
 46 aatctccttc aaggtgtgga tgtgtggggg ccgcatggag gacatcccct gctccaggg 1200
 47 gggccatata tacaggaagt atgtgccta caaggtcccg gccggagtca gcctggccc 1260
 48 gaaccttaag cgggtggcgg aagtgtggat ggatgagtac gcagagtaca tttaccagcg 1320
 49 ccggcctgaa taccgccacc tctccgctgg ggatgtcgca gtccagaaaa agctccgcag 1380
 50 ctcccttaac tgcaagagtt tcaagtgggt tatgacgaag atagcctggg acctgcccaa 1440
 51 attctaccca ccggtggagc ccccggtctc agcttggggg gagatccgaa atgtgggcac 1500
 52 agggctgtgt gcagacacaa agcacggggc cttgggctcc ccactaaggc tagagggctg 1560
 53 cgtccgaggc cgtggggagg ctgcctggaa caacatgcag gtattcacct tcacctggag 1620
 54 agaggacatc cggcctggag acccccagca caccaagaag ttctgctttg atgccatttc 1680
 55 ccacaccagc cctgtcacgc tgtacgactg ccacagcatg aagggaacc agctgtggaa 1740
 56 ataccgcaaa gacaagacct tgtaccacc tgtcagtggc agctgcatgg actgcagtga 1800
 57 aagtgaccat aggatcttca tgaacacctg caacccatcc tctctcacc agcagtggct 1860

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/001,851

DATE: 12/11/2001

TIME: 11:57:51

Input Set : A:\10147561.app

Output Set: N:\CRF3\12112001\I001851.raw

```

58 gtttgaacac accaactcaa cagtcttgga aaaattcaat aggaactgag ccctcatgtc 1920
59 cccttggcag gccccccagg gtctggcact cactgcagac ttctctttc aagggaggca 1980
60 gggcccctgt gggcactagg tgtaaaaggt gctggccaaa tggttcaggg tgaagagggc 2040
61 tcttgattca ggggctgggg tctgcctggt ccttgagccc ctgagttgtg ggggtagggt 2100
62 gaagagcata tcccacaaga ggccccacag ggagcagaga ctgctttaat ccctgctgac 2160
63 atcacggaag agcaacagag ccttttcaac ttgtcacta tgtccccttg aacattatgt 2220
64 ggagaacac caaggtagcc taggccaccc aaaagtgagt cctgcgaggt tgcccagccc 2280
65 tcagatggct ctctacatg atggtgcttt agaaacaaag gtaaaatttg cctgtttggg 2340
66 gcagctttta gtatcgatgc cactcatctg cagcagaaga gaaagaagtc ctcttggggc 2400
67 tttttagttt ctgccgtcct ggggggaaca ttgcagttac tgcacagctt ctgttctctg 2460
68 tcacaacccc aggtgatttg gtccgggtcaa aggccatact tggggcccta agagtgttca 2520
69 gtattgaatg ctgatcagct gccagggtgag gagtcaagag agggagcccc cctagacatt 2580
70 tctttgcagc tatggacatg cgggatattc cccctgctc tctgggtatt tgaaatgtca 2640
71 attttagcac tctccaggca caaggacagc ccagcaccag ctttacaggg cagtgtttca 2700
72 gatggccctg agcccacgga aaaggccagg tagacctcca aactagaaat gctggctgat 2760
73 ttgcctgat ccatgcttcc atttccctgt ctctcttccc caggcaatta ctggcctcaa 2820
74 aagaggaaca gaggtgctgc gaggtgctca cctcacagag tctggaggcc tccaggatca 2880
75 actgtgggca aagtgcctgc ctctgacctc atcatggttc tagttctcat acagaactcc 2940
76 agaattttta aagaactcta taattggatt gcaaaactag atgctacata ggattctggt 3000
77 attccacatc caatatggat ttctagaatg ctgtgattaa aggagccagc cagggtgtaat 3060
78 acagtcaagg cagccccag cctagagaca atctgtgaaa tccaaagttg gtggtgttgg 3120
79 gaaagcaggg ggacatgtgt ccctcagctc agcagaggct gtggtacaac atggtccttg 3180
80 gtgaagacct gcacccctgg aacctcccac catcatcaca actgtagtct catttgagct 3240
81 ggagaaaaga acccgacgtc ccacagccag atatacacc agctccatgc cagcccttca 3300
82 tgtttacctt ttgctttggt aattacatgt cagactccta gagggcctcc agactaatag 3360
83 gaagcatttc tgtaaccaac ctgccaccca ctgattcaga aatggaaatc acattccaca 3420
84 atctatggct tccaccagct agcccaggaa atacttgaaa tcagcattcc aattagtgtt 3480
85 gagtctcttg attgtgtcat ttaccaatta aataactgag acctaagtct gggaacagag 3540
86 ccagcaatct gcctttgaga tgctggcaga tctcaaggcc atcaattatt gggggaggga 3600
87 gggacaaaaca ctcccaatca tccaccagtc agactgaatg tgtagctggc gaggaattac 3660
88 ttccacttct ggcccagcac aagccctgct ttggccacct gtctgcaaga gaggcggccc 3720
89 ctgtgcttgc aacgcttacg tgttgatccc agtgccttt tccaaatgag tgctgtagct 3780
90 ttagaagtgg ccctctatag aaagaagtca aaagatgagg ccccttctag aatctaggat 3840
91 aacaagagtg ttgacagttt gaggagtcca attgagattc atcatcaaag agcaatgcag 3900
92 cgtcgtaaaa ataaaaactg tgccctttta aaagaaaaat gcaaatatag agcaaatccc 3960
93 taaacttgaa aaaaaaaaaa aaaaa 3985

```

96 <210> SEQ ID NO: 2

97 <211> LENGTH: 603

98 <212> TYPE: PRT

99 <213> ORGANISM: Homo sapiens

101 <400> SEQUENCE: 2

```

102 Met Arg Arg Lys Glu Lys Arg Leu Leu Gln Ala Val Ala Leu Val Leu
103   1           5           10           15
105 Ala Ala Leu Val Leu Leu Pro Asn Val Gly Leu Trp Ala Leu Tyr Arg
106           20           25           30
108 Glu Arg Gln Pro Asp Gly Thr Pro Gly Gly Ser Gly Ala Ala Val Ala
109           35           40           45
111 Pro Ala Ala Gly Gln Gly Ser His Ser Arg Gln Lys Lys Thr Phe Phe
112   50           55           60

```

RAW SEQUENCE LISTING

DATE: 12/11/2001

PATENT APPLICATION: US/10/001,851

TIME: 11:57:51

Input Set : A:\10147561.app

Output Set: N:\CRF3\12112001\I001851.raw

```

114 Leu Gly Asp Gly Gln Lys Leu Lys Asp Trp His Asp Lys Glu Ala Ile
115 65 70 75 80
117 Arg Arg Asp Ala Gln Arg Val Gly Asn Gly Glu Gln Gly Arg Pro Tyr
118 85 90 95
120 Pro Met Thr Asp Ala Glu Arg Val Asp Gln Ala Tyr Arg Glu Asn Gly
121 100 105 110
123 Phe Asn Ile Tyr Val Ser Asp Lys Ile Ser Leu Asn Arg Ser Leu Pro
124 115 120 125
126 Asp Ile Arg His Pro Asn Cys Asn Ser Lys Arg Tyr Leu Glu Thr Leu
127 130 135 140
129 Pro Asn Thr Ser Ile Ile Ile Pro Phe His Asn Glu Gly Trp Ser Ser
130 145 150 155 160
132 Leu Leu Arg Thr Val His Ser Val Leu Asn Arg Ser Pro Pro Glu Leu
133 165 170 175
135 Val Ala Glu Ile Val Leu Val Asp Asp Phe Ser Asp Arg Glu His Leu
136 180 185 190
138 Lys Lys Pro Leu Glu Asp Tyr Met Ala Leu Phe Pro Ser Val Arg Ile
139 195 200 205
141 Leu Arg Thr Lys Lys Arg Glu Gly Leu Ile Arg Thr Arg Met Leu Gly
142 210 215 220
144 Ala Ser Val Ala Thr Gly Asp Val Ile Thr Phe Leu Asp Ser His Cys
145 225 230 235 240
147 Glu Ala Asn Val Asn Trp Leu Pro Pro Leu Leu Asp Arg Ile Ala Arg
148 245 250 255
150 Asn Arg Lys Thr Ile Val Cys Pro Met Ile Asp Val Ile Asp His Asp
151 260 265 270
153 Asp Phe Arg Tyr Glu Thr Gln Ala Gly Asp Ala Met Arg Gly Ala Phe
154 275 280 285
156 Asp Trp Glu Met Tyr Tyr Lys Arg Ile Pro Ile Pro Pro Glu Leu Gln
157 290 295 300
159 Lys Ala Asp Pro Ser Asp Pro Phe Glu Ser Pro Val Met Ala Gly Gly
160 305 310 315 320
162 Leu Phe Ala Val Asp Arg Lys Trp Phe Trp Glu Leu Gly Gly Tyr Asp
163 325 330 335
165 Pro Gly Leu Glu Ile Trp Gly Gly Glu Gln Tyr Glu Ile Ser Phe Lys
166 340 345 350
168 Val Trp Met Cys Gly Gly Arg Met Glu Asp Ile Pro Cys Ser Arg Val
169 355 360 365
171 Gly His Ile Tyr Arg Lys Tyr Val Pro Tyr Lys Val Pro Ala Gly Val
172 370 375 380
174 Ser Leu Ala Arg Asn Leu Lys Arg Val Ala Glu Val Trp Met Asp Glu
175 385 390 395 400
177 Tyr Ala Glu Tyr Ile Tyr Gln Arg Arg Pro Glu Tyr Arg His Leu Ser
178 405 410 415
180 Ala Gly Asp Val Ala Val Gln Lys Lys Leu Arg Ser Ser Leu Asn Cys
181 420 425 430
183 Lys Ser Phe Lys Trp Phe Met Thr Lys Ile Ala Trp Asp Leu Pro Lys
184 435 440 445
186 Phe Tyr Pro Pro Val Glu Pro Pro Ala Ala Ala Trp Gly Glu Ile Arg

```

RAW SEQUENCE LISTING

DATE: 12/11/2001

PATENT APPLICATION: US/10/001,851

TIME: 11:57:51

Input Set : A:\10147561.app

Output Set: N:\CRF3\12112001\I001851.raw

```

187      450      455      460
189 Asn Val Gly Thr Gly Leu Cys Ala Asp Thr Lys His Gly Ala Leu Gly
190 465      470      475      480
192 Ser Pro Leu Arg Leu Glu Gly Cys Val Arg Gly Arg Gly Glu Ala Ala
193      485      490      495
195 Trp Asn Asn Met Gln Val Phe Thr Phe Thr Trp Arg Glu Asp Ile Arg
196      500      505      510
198 Pro Gly Asp Pro Gln His Thr Lys Lys Phe Cys Phe Asp Ala Ile Ser
199      515      520      525
201 His Thr Ser Pro Val Thr Leu Tyr Asp Cys His Ser Met Lys Gly Asn
202      530      535      540
204 Gln Leu Trp Lys Tyr Arg Lys Asp Lys Thr Leu Tyr His Pro Val Ser
205 545      550      555      560
207 Gly Ser Cys Met Asp Cys Ser Glu Ser Asp His Arg Ile Phe Met Asn
208      565      570      575
210 Thr Cys Asn Pro Ser Ser Leu Thr Gln Gln Trp Leu Phe Glu His Thr
211      580      585      590
213 Asn Ser Thr Val Leu Glu Lys Phe Asn Arg Asn
214      595      600
217 <210> SEQ ID NO: 3
218 <211> LENGTH: 1809
219 <212> TYPE: DNA
220 <213> ORGANISM: Homo sapiens
222 <400> SEQUENCE: 3
223 atgaggcgga aggagaagcg gctcctgcag gcggtggcgc tgggtgctggc ggccctggtc 60
224 ctccctgccc acgtggggct ttgggcgctg tacgcgagc gccagcccga cggcaccctc 120
225 ggggggatcgg gggcggcgggt ggcgcggcgc gcgggacagg gctcacacag tcgacaaaag 180
226 aaaacgtttt tcttgggaga tgggcagaag ctgaaggact ggcatgacaa ggaggccatc 240
227 cggagggacg ctcagcgcgt aggaaatgga gaacaaggaa gacctaccc catgaccgat 300
228 gctgagagag tggatcaggc ataccgagaa aatggattta acatctacgt cagtataaaa 360
229 atctccttga atcgctctct ccagatatc cggcacccaa actgcaacag caagcgctac 420
230 ctggagacac ttcccaacac aagcatcatc atccccttcc acaacgaggg ctggtcctcc 480
231 ctccctccga cgtccacag tgtgtcaat cgctgcctc cagagctggc cgcgagatt 540
232 gtactggtcg acgacttcag tgatcgagag cacctgaaga agcctcttga agactacatg 600
233 gcccttttcc ccagtgtgag gattcttoga accaagaaac gggaagggct gataaggacc 660
234 cgaatgctgg gggcctcagt ggcaactggg gatgtcatca cattcttga ttacactgt 720
235 gaagccaatg tcaactggct tcccccttg cttgaccgca ttgctcgga ccgcaagacc 780
236 attgtgtgcc cgatgattga tgtaattgac catgacgact ttcggtacga gacacaggca 840
237 ggggatgcca tgcggggagc ctttgactgg gagatgtact acaagcggat cccgatccct 900
238 ccagaactgc agaaagctga cccagcgac ccatttgagt ctcccgtgat ggccggtgga 960
239 ctgttcgccc tggatcgga gtggttctgg gaactcggcg ggtatgacc aggcttgga 1020
240 atctggggag gggagcagta tgaaatctcc ttcaagggtg ggatgtgtgg gggccgcatg 1080
241 gaggacatcc cctgctccag ggtgggccat atctacagga agtatgtgcc ctacaaggtc 1140
242 ccggccggag tcagcctggc ccggaacctt aagcgggtgg ccgaagtgtg gatggatgag 1200
243 tacgcagagt acatttacca gcgccggcct gaataccgcc acctctccgc tggggatgtc 1260
244 gcagtccaga aaaagctccg cagctccctt aactgcaaga gtttcaagtg gtttatgacg 1320
245 aagatagcct gggacctgcc caaattctac ccaccgtgg agccccggc tgcagcttgg 1380
246 ggggagatcc gaaatgtggg cacagggtcg tgtgcagaca caaagcacgg ggccttgggc 1440
247 tccccactaa ggctagagg ctgcgtccga ggccgtgggg aggctgcctg gaacaacatg 1500

```

RAW SEQUENCE LISTING

DATE: 12/11/2001

PATENT APPLICATION: US/10/001,851

TIME: 11:57:51

Input Set : A:\10147561.app

Output Set : N:\CRF3\12112001\I001851.raw

```

248 caggtattca ccttcacctg gagagaggac atccggcctg gagaccccca gcacaccaag 1560
249 aagttctgct ttgatgccat ttcccacacc agccctgtca cgctgtacga ctgccacagc 1620
250 atgaagggca accagctgtg gaaataccgc aaagacaaga ccctgtacca ccctgtcagt 1680
251 ggcagctgca tggactgcag tgaaagtgc cataggatct tcatgaacac ctgcaaccca 1740
252 tctctctca cccagcagtg gctgtttgaa cacaccaact caacagtctt ggaaaaattc 1800
253 aataggaaac 1809
256 <210> SEQ ID NO: 4
W--> 257 <400> SEQUENCE: 4
W--> 258 000
261 <210> SEQ ID NO: 5
W--> 262 <400> SEQUENCE: 5
W--> 263 000
266 <210> SEQ ID NO: 6
W--> 267 <400> SEQUENCE: 6
W--> 268 000
271 <210> SEQ ID NO: 7
W--> 272 <400> SEQUENCE: 7
W--> 273 000
276 <210> SEQ ID NO: 8
W--> 277 <400> SEQUENCE: 8
W--> 278 000
281 <210> SEQ ID NO: 9
W--> 282 <400> SEQUENCE: 9
W--> 283 000
286 <210> SEQ ID NO: 10
W--> 287 <400> SEQUENCE: 10
W--> 288 000
291 <210> SEQ ID NO: 11
292 <211> LENGTH: 4158
293 <212> TYPE: DNA
294 <213> ORGANISM: Homo sapiens
296 <400> SEQUENCE: 11
297 tcggcggaag atggcggccg gcgaaaggag ctggtgcctg tgcaagttgt tgagggtttt 60
298 ttattcatta ttcttcctg ggctcattgt atgtggaact ttatgtgtgt gtttggtcat 120
299 tgtccttttg ggaatcagac tgctgtctaca gagaaagaaa aaattagtgt caactagcaa 180
300 aaatgggaaa aatcaaatgg tgattgcatt ttttcatcca tactgcaatg ctggtggagg 240
301 aggagaaaaga gttttatggt gtgctttaag agccctgcag aaaaagtatc ctgaagcagt 300
302 ttatgttgtt tataccggcg atgttaatgt caacggtcaa cagatactag aagggtgctt 360
303 cagaagattt aacatcagat taattcacc agtgcagttt gtttttttaa ggaaacgcta 420
304 tcttgtggaa gattcactgt atcctcactt cacactgctg ggccaaagtc taggatccat 480
305 ttttcttggc tgggaagctc taatgcagtg tgttcctgat gtttacattg attcaatggg 540
306 atacgctttt acgcttcctc tgtttaagta tatagggggt tgccaagttg gaagctatgt 600
307 tcattatcct actatcagca ccgacatgct ctctgtagtg aagaatcaaa atattggatt 660
308 taataatgca gccttcatta ccaggaatcc ttttctcagc aaagtaaagc tcatctacta 720
309 ctatttattt gcttttattt atggacttgt tggttcttgc agtgatgtag tcatggtcaa 780
310 ttcttcttgg acactaaacc atattctctc actatggaaa gttgggaatt gcaactaacat 840
311 tgtttatoca ctttgtgatg tgcagacatt tctggacatt cccttacatg agaaaaagat 900
312 gacccagga catttgctgg tttctgttgg ccagtttagg ccggaaaaga atcatccatt 960
313 gcagatcaga gccttgcta aattgctgaa taagaagatg gttgagtcac ctcttcgct 1020

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/001,851

DATE: 12/11/2001

TIME: 11:57:52

Input Set : A:\10147561.app

Output Set: N:\CRF3\12112001\I001851.raw

L:11 M:270 C: Current Application Number differs, Replaced Current Application Number
L:257 M:283 W: Missing Blank Line separator, <400> field identifier
L:258 M:300 W: (50) Intentionally skipped Sequence, : Sequence Id (4) SEQUENCE:
L:262 M:283 W: Missing Blank Line separator, <400> field identifier
L:263 M:300 W: (50) Intentionally skipped Sequence, : Sequence Id (5) SEQUENCE:
L:267 M:283 W: Missing Blank Line separator, <400> field identifier
L:268 M:300 W: (50) Intentionally skipped Sequence, : Sequence Id (6) SEQUENCE:
L:272 M:283 W: Missing Blank Line separator, <400> field identifier
L:273 M:300 W: (50) Intentionally skipped Sequence, : Sequence Id (7) SEQUENCE:
L:277 M:283 W: Missing Blank Line separator, <400> field identifier
L:278 M:300 W: (50) Intentionally skipped Sequence, : Sequence Id (8) SEQUENCE:
L:282 M:283 W: Missing Blank Line separator, <400> field identifier
L:283 M:300 W: (50) Intentionally skipped Sequence, : Sequence Id (9) SEQUENCE:
L:287 M:283 W: Missing Blank Line separator, <400> field identifier
L:288 M:300 W: (50) Intentionally skipped Sequence, : Sequence Id (10) SEQUENCE:
L:529 M:283 W: Missing Blank Line separator, <400> field identifier
L:530 M:300 W: (50) Intentionally skipped Sequence, : Sequence Id (14) SEQUENCE:
L:534 M:283 W: Missing Blank Line separator, <400> field identifier
L:535 M:300 W: (50) Intentionally skipped Sequence, : Sequence Id (15) SEQUENCE:
L:539 M:283 W: Missing Blank Line separator, <400> field identifier
L:540 M:300 W: (50) Intentionally skipped Sequence, : Sequence Id (16) SEQUENCE:
L:544 M:283 W: Missing Blank Line separator, <400> field identifier
L:545 M:300 W: (50) Intentionally skipped Sequence, : Sequence Id (17) SEQUENCE:
L:549 M:283 W: Missing Blank Line separator, <400> field identifier
L:550 M:300 W: (50) Intentionally skipped Sequence, : Sequence Id (18) SEQUENCE:
L:554 M:283 W: Missing Blank Line separator, <400> field identifier
L:555 M:300 W: (50) Intentionally skipped Sequence, : Sequence Id (19) SEQUENCE: